Model-Based Curve Evolution Technique for Image Segmentation *

Andy Tsai [1] Anthony Yezzi, Jr.[2] William Wells III [3,4] Clare Tempany [4]
Dewey Tucker [1] Ayres Fan [1] W. Eric Grimson [1,3] Alan Willsky [1]

[1] Dept. of EECS, Massachusetts Institute of Technology, Cambridge, MA
[2] School of ECE, Georgia Institute of Technology, Atlanta, GA
[3] Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA
[4] Brigham and Women's Hospital, Harvard Medical School, Boston, MA

Abstract

We propose a model-based curve evolution technique for segmentation of images containing known object types. In particular, motivated by the work of Leventon, Grimson, and Faugeras [4], we derive a parametric model for an implicit representation of the segmenting curve by applying principal component analysis to a collection of signed distance representations of the training data. The parameters of this representation are then calculated to minimize an objective function for segmentation. We found the resulting algorithm to be computationally efficient, able to handle multidimensional data, robust to noise and initial contour placements, while at the same time, avoiding the need for point correspondences during the training phase of the algorithm. We demonstrate this technique by applying it to two medical applications.

1 Introduction

Image segmentation algorithms often face difficult challenges such as poor image contrast, noise, and missing or diffuse boundaries. In these settings, without a prior model to constrain the segmentation, the algorithms often fail, mostly due to the under-determined nature of the segmentation process. Model-based segmentation algorithms can often alleviate this problem. We propose such an algorithm in this paper.

Our work shares common aspects with a number of model-based image segmentation algorithms in the literature. In particular, Cootes *et al.* [3] developed a parametric point distribution model for describing the segmenting curve using linear combinations of the eigenvectors that reflect the shape variations in the training database. The shape and pose parameters of this point model are determined to match the strongest image gradients around each landmark point. Staib and Duncan [7] introduced a parametric point model based on an elliptic Fourier decomposition of the segmenting curve. The parameters of their curve are calculated to optimize the match between the segmenting curve and the gradient of the image. Chakraborty *et al.* [1] then extended this approach to a hybrid segmentation model that incorporates both gradient and region-homogeneity information. Recently, Wang and Staib [8] developed a statistical point model for the segmenting curve by applying principal component analysis on the covariance matrices of point locations derived from the training set. This approach finds the boundaries of an object while simultaneously attempts to address the point correspondence problem. Leventon *et al.* [4] proposed a less restrictive model-based segmenter. In particular, they incorporated shape information as a prior model to restrict the flow of the geodesic active contour. Their prior parametric shape model is derived by performing principal component analysis on a collection of signed distance maps of the training shape. The segmenting curve then evolves according to two competing forces: (1) the gradient force of the image, and (2) the force exerted by the estimated shape with the parameters of the shape calculated based on the image gradient and the current position of the curve.

Our work is also closely related to many region-based active contour models in the literature, in particular [2] and [9]. In general, these region-based models enjoy a number of attractive properties over gradient-based techniques for segmentation including greater robustness to noise (by avoiding derivatives of the image intensity) and initial contour placement (by being less local than most edge-based approaches).

In our algorithm, we adopt the implicit parametric representation of the segmenting curve proposed in [4], and calculate the parameters of this implicit model to minimize the region-based energy functional proposed in [2] for image segmentation. The resulting algorithm is found to be computationally efficient and robust to noise (since the evolution of the curve has reduced

*This work was supported by ONR grant N00014-00-1-0089, AFOSR grant F49620-98-1-0349, NSF ERC grant under Johns Hopkins Agreement 8810274, NIH 1P41RR13218, and NIH R01 grant AG 19513-01.

0-7695-1272-0/01 $10.00 © 2001 IEEE

degrees of freedom), has an extended capture range (since the segmentation functional is region-based instead of edge-based), and does not require point correspondences (since the curve representation is set in an Eulerian framework). Though in this paper we focus on the region-based segmentation model of [2], it is important to point out that our algorithmic framework is equally applicable to other region-based models.

The rest of this paper is organized as follows. Section 2 describes an approach to align all the training shapes in the database to eliminate variations in the shapes due to pose differences. Based on this aligned training set, we show in Section 3 the development of an implicit parametric representation of the segmenting curve. Section 4 describes our algorithmic framework which involves optimizing the pose and shape parameters of our curve model over a region-based energy functional for image segmentation. In Section 5, we show the application of this technique to two medical applications. We conclude in Section 6 with a summary of the paper.

2 Binary Shape Alignment

Let the training set $\mathcal{T}$ consist of a set of n binary images $\{I^1, I^2, ..., I^n\}$, each with values of one inside and zero outside the object. The idea is to calculate the set of pose parameters $\{\mathbf{p}^1, \mathbf{p}^2, ..., \mathbf{p}^n\}$ used to transform the n binary images to jointly align them. In 2D, $\mathbf{p}$ consists of x-, y-translation, scale, and rotation. In 3D, $\mathbf{p}$ consists of x-, y-, and z-translation, scale, and the three axis rotations yaw, pitch, and roll. In 2D, the transformed image of I^i based on the pose parameter $\mathbf{p}^i$ is defined as

$$\boldsymbol{T}^i = I^i(\, T(\mathbf{p}^i) \begin{bmatrix} x & y & 1 \end{bmatrix}^T)$$

where $T(\mathbf{p}^i)$ is a transformation matrix parametrized by $\mathbf{p}^i$ that transforms the coordinate of one image (x, y) into the coordinate of another image. The 3D version of $\boldsymbol{T}^i$ is defined in a similar manner.

An effective strategy to align the n binary images is to descend along the following energy functional:

$$E_{align} = \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \left\{ \frac{\int_\Omega (\boldsymbol{T}^i - \boldsymbol{T}^j)^2 \, dA}{\int_\Omega (\boldsymbol{T}^i + \boldsymbol{T}^j)^2 \, dA} \right\} \quad (1)$$

where Ω denotes the image domain. Minimizing (1) is equivalent to simultaneously minimizing the difference between any pair of binary images in the training database. The area normalization term in the denominator of (1) is employed to protect against all the images from shrinking to improve the cost function.

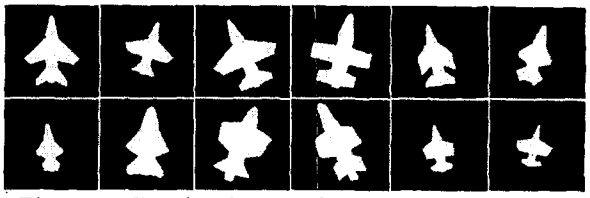

Figure 1: Training data: twelve binary shape models.

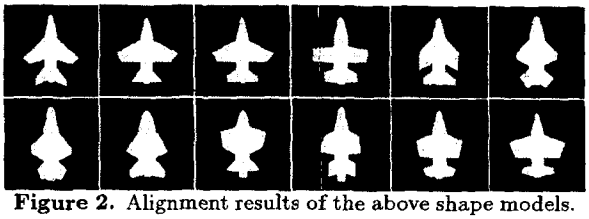

Figure 2. Alignment results of the above shape models.

The gradient of E_{align}, taken with respect to $\mathbf{p}_i$ for any i, is given by

$$\nabla_{\mathbf{p}^i} E_{align} = \sum_{\substack{j=1 \\ j \neq i}}^{n} \left\{ \frac{2\int_\Omega (\boldsymbol{T}^i - \boldsymbol{T}^j) \nabla_{\mathbf{p}^i} \boldsymbol{T}^i dA}{\int_\Omega (\boldsymbol{T}^i + \boldsymbol{T}^j)^2 \, dA} - \frac{2\int_\Omega (\boldsymbol{T}^i - \boldsymbol{T}^j)^2 \, dA \int_\Omega (\boldsymbol{T}^i + \boldsymbol{T}^j) \nabla_{\mathbf{p}^i} \boldsymbol{T}^i dA}{\left(\int_\Omega (\boldsymbol{T}^i + \boldsymbol{T}^j)^2 \, dA\right)^2} \right\}$$

where $\nabla_{\mathbf{p}^i} \boldsymbol{T}^i$ is the gradient of the transformed image $\boldsymbol{T}^i$ taken with respect to $\mathbf{p}^i$. Since this alignment problem is under-determined, we artificially fix the pose of one of the shapes, and calculate the remaining pose parameters using the above described gradient descent approach. To illustrate, a training set consisting of twelve synthetic binary fighter jets is shown in Figure 1. In this example, the pose parameter associated with the fighter jet in the upper-left corner of the figure is chosen to be fix. The aligned version of this data set is shown in Figure 2.

3 Parametric Shape Model

Instead of using a point model to represent the curve [1, 3, 7, 8], we follow the lead of [4] and [6] by choosing the signed distance function as our representation. In particular, the boundaries of each of the n shapes in the database are embedded as the zero level set of n separate signed distance functions $\{\Psi_1, \Psi_2, ..., \Psi_n\}$ with negative distances assigned to the inside and positive distances assigned to the outside of the object. Using the technique developed in [4], we compute $\bar{\phi}$, the mean level set function of the shape database, as the average of these n signed dis-

Figure 3. Shape variability of the fighter jet. (a) The mean shape. (b) & (c) $\pm 1\sigma$ variance of the 1^{st} principal mode. (d) & (e) $\pm 1\sigma$ variance of the 2^{nd} principal mode.

tance functions, i.e. $\bar{\phi} = \frac{1}{n}\sum \Psi_i$. To extract the shape variabilities, $\bar{\phi}$ is subtracted from each of the n signed distance functions to create n mean-offset functions $\{\tilde{\Psi}_1, \tilde{\Psi}_2, ..., \tilde{\Psi}_n\}$. These mean-offset functions are then used to capture the variabilities of a shape. In particular, each function $\tilde{\Psi}_i$, consisting of N pixels, is stacked in lexicographical order to form a column vector in the $N \times n$-dimensional shape-variability matrix $\mathcal{S}$. Eigenvalue decomposition is employed to factor $\frac{1}{n}\mathcal{S}\mathcal{S}^T$ as:

$$\frac{1}{n}\mathcal{S}\mathcal{S}^T = U\Sigma U^T$$

where U is an $N \times n$ matrix whose columns represent the n orthogonal modes of variation in the shape, and Σ is an $n \times n$ diagonal matrix whose diagonal elements represent the corresponding eigenvalues. Each column in U is then reordered into the dimension of the data set to yield a maximum of n different principal modes or eigenshapes $\{\phi_1, \phi_2, ..., \phi_n\}$.

Let k, which is selected prior to segmentation, be the number of modes to consider. Shape can now be represented as the zero level set of the function defined by the linear weighted sum of the k eigenshapes plus its mean shape, i.e.

$$\phi(\mathbf{w}) = \bar{\phi} + \sum_{i=1}^{k} w_i\, \phi_i \qquad (2)$$

where $\mathbf{w} = \{w_1, w_2, ..., w_k\}$ are the weights for the k eigenshapes with the variances of these weights $\{\sigma_1^2, \sigma_2^2, ..., \sigma_k^2\}$ given by the eigenvalues calculated earlier. To illustrate the capabilities of this parametric shape encoding scheme, we show in Figure 3 the mean shape of the fighter jet as well as its shape variances based on varying its first two principal modes by $\pm 1\sigma$.

To have the flexibility of capturing pose variabilities, pose parameter $\mathbf{p}$ is added as another parameter to the level set function of (2). With this addition, the implicit description of shape or, in our segmentation framework, the segmenting curve, is given by the zero level set of the following level set function:

$$\phi(\mathbf{w}, \mathbf{p}) = \bar{\phi}(\mathbf{p}) + \sum_{i=1}^{k} w_i\, \phi_i(\mathbf{p}) \qquad (3)$$

where each ϕ_i as well as $\bar{\phi}$ are now a function of $\mathbf{p}$.

4 Statistical Model for Segmentation

Assume that the domain of the observed image I is formed by two regions of distinct intensities. In addition, let R^u and R^v denote the regions inside and outside the segmenting curve, respectively. Chan and Vese [2] proposed the following energy functional for segmenting I:

$$\begin{aligned} E_{cv} &= -\left(\mu^2 A_u + \nu^2 A_v\right) \\ &= -\left(\frac{S_u^2}{A_u} + \frac{S_v^2}{A_v}\right) \end{aligned} \qquad (4)$$

where μ, S_u, A_u, ν, S_v, and A_v represent the mean intensity, the sum intensity, and the area in R^u and R^v, respectively. This functional can be viewed as a piecewise constant generalization of the Mumford-Shah segmentation functional [5]. This reduced form uses the area-weighted mean intensities inside and outside the curve as a measure of how well the segmenting curve has separated the two regions in I. In our framework, the parameters S_u, A_u, S_v, and A_v in (4) can be written in terms of $\phi(\mathbf{w}, \mathbf{p})$ as: $S_u = \int_\Omega I\mathcal{H}(\phi(\mathbf{w}, \mathbf{p}))\, dA$, $A_u = \int_\Omega \mathcal{H}(\phi(\mathbf{w}, \mathbf{p}))\, dA$, $S_v = \int_\Omega I\mathcal{H}(-\phi(\mathbf{w}, \mathbf{p}))\, dA$, and $A_v = \int_\Omega \mathcal{H}(-\phi(\mathbf{w}, \mathbf{p}))\, dA$, where the Heaviside function

$$\mathcal{H}(\phi(\mathbf{w}, \mathbf{p})) = \begin{cases} 1 & \text{if } \phi(\mathbf{w}, \mathbf{p}) < 0 \\ 0 & \text{if } \phi(\mathbf{w}, \mathbf{p}) \geq 0 \end{cases}.$$

Gradient descent is employed to minimize (4). The gradient of E_{cv}, taken with respect to $\mathbf{w}$ and $\mathbf{p}$, is given by:

$$\begin{aligned} \nabla_{(\mathbf{w},\mathbf{p})} E_{cv} &= -2\left(\mu \nabla_{(\mathbf{w},\mathbf{p})} S_u + \nu \nabla_{(\mathbf{w},\mathbf{p})} S_v\right) \\ &\quad + \left(\mu^2 \nabla_{(\mathbf{w},\mathbf{p})} A_u + \nu^2 \nabla_{(\mathbf{w},\mathbf{p})} A_v\right) \end{aligned}$$

with

$$\begin{aligned} \nabla_{w_i} A_u &= -\nabla_{w_i} A_v &&= -\int_{\vec{C}} \phi_i(\mathbf{p})\, ds \\ \nabla_{w_i} S_u &= -\nabla_{w_i} S_v &&= -\int_{\vec{C}} I\, \phi_i(\mathbf{p})\, ds \\ \nabla_{p_i} A_u &= -\nabla_{p_i} A_v &&= -\int_{\vec{C}} \nabla_{p_i}\phi(\mathbf{w}, \mathbf{p})\, ds \\ \nabla_{p_i} S_u &= -\nabla_{p_i} S_v &&= -\int_{\vec{C}} I\, \nabla_{p_i}\phi(\mathbf{w}, \mathbf{p})\, ds \end{aligned}$$

where the segmenting curve $\vec{C}$ is given by the zero level set of $\phi(\mathbf{w}, \mathbf{p})$, and $\nabla_{p_i}\phi(\mathbf{w}, \mathbf{p})$ is the gradient of $\phi(\mathbf{w}, \mathbf{p})$ taken with respect to the ith component of $\mathbf{p}$. This gradient descent approach is used to calculate both parameters $\mathbf{w}$ and $\mathbf{p}$ for image segmentation.

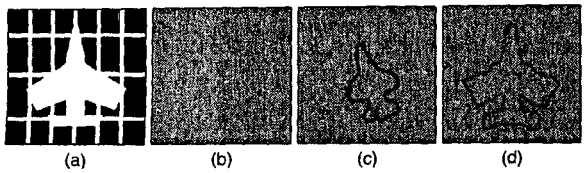

Figure 4. Segmentation of a noisy fighter jet with missing edges. (a) Original binary image surrounded by line clutter. (b) Image in (a) with additive Gaussian noise. (c) Blue curve shows the initializing contour. (d) Red curve shows the final contour.

To illustrate the use of this model for segmentation, we show in Figure 4(a) a fighter jet surrounded by horizontal and vertical line clutters. The presence of these lines creates missing edges in the fighter jet which can cause problems in conventional segmentation algorithms that do not rely on prior shape information. Figure 4(b) shows this same image contaminated by additive Gaussian noise. Knowing that the object in the image is a fighter jet, we employ the pose-aligned database shown in Figure 2 to derive an implicit parametric curve model for the fighter jet in the form of (3). In this example, we choose $k = 6$. The zero level set of $\bar{\phi}$ is employed as the initial shape of the starting curve. Figure 4(c) shows the position of this curve. The parameters of the segmenting curve, $\mathbf{w}$ and $\mathbf{p}$, are calculated to minimizing the energy functional described in (4). Figure 4(d) shows the final shape and position of the segmenting curve.

5 Applications to Medical Imagery

We applied our algorithm to two medical applications. Section 5.1 illustrates a 2D example, while Section 5.2 illustrates a 3D example.

5.1 Left Ventricle Segmentation of Cardiac MRI

Cardiac MRI is an important clinical tool used to provide 4D (temporal as well as spatial) information about the heart. Typically, one study generates 80-120 2D images of a single patient's heart. In a variety of clinical scenarios (such as assessing cardiac function and diagnosing cardiac diseases), it is important to extract the boundaries of the left ventricle from this data set. Manual tracing of the left ventricle from such a large data sets is both tedious and time-consuming.

Conventional automated segmentation techniques usually encounter difficulties in segmenting the left ventricle because 1) the intensity contrast between the ventricle and the myocardium is low (due to the smearing of the blood pool in the ventricle into the myocardium), and 2) the boundaries of the left ventricle are missing at certain locations due to the presence of protruding papillary muscles which have the same intensity profile as the myocardium.

In our experiment, we equally divided the 100 2D images from a single patient into two sets: a training set and a test set. Fifty 4D interactive segmentations of the left ventricle from the training set form the shape database. Part of this database is shown in Figure 5. Figure 6 shows the mean shape of the left ventricle and its shape variances based on varying the first four eigenshapes by $\pm 1\sigma$. Figure 7 shows part of the segmentation result of the testing set (red curves) with $k = 25$. These results are comparable to the ones given by the 4D interactive cardiac MRI segmenter (green curves).

5.2 Prostate Segmentation of Pelvic MRI Taken With Endorectal Coil

Pelvic MRI, when taken in conjunction with an endorectal coil (ERC)–a receive-only surface coil placed within the rectum–provides high resolution images of the prostate which are commonly used for staging of prostate cancer. For assignment of MR-guided radiation therapy, the segmentation of the prostate gland from these images is required. Manual outlining of sequential 2D cross-sectional slices of the prostate is tedious and difficult. Automatic segmentation of the prostate is also difficult because the prostate is a small glandular structure buried deep within the pelvic region and surrounded by a variety of different tissues which show up as varying intensity levels on MRI. This segmentation problem is further complicated by an artifact caused by the use of ERC called the near-field effect. This effect causes an intensity artifact to appear in the tissues surrounding the ERC. This artifact–seen as a white circular halo surrounding the rectum in each image slice of Figure 10–bleaches out the anterior borders of the prostate, making automatic prostate segmentation problem even more difficult.

In Figure 8, we show the training data which consists of eight 3D binary shape models of the prostate gland–obtained by stacking together 2D expert hand segmentations of eight patients' pelvic MRIs taken with an ERC. After alignment but prior to shape training, the 3D shape models are smoothed to remove the "step-like" artifact along the axial direction of the prostate. Figure 9 shows the mean shape of the prostate gland as well as its shape variances based on varying the first four eigenshapes by $\pm 1\sigma$.

To accentuate the boundaries of the prostate gland as well as to minimize the intensity artifact caused by the ERC, the pelvic MRI data set I_{MRI} is transformed to a bimodal data set I by applying the following map:

$$I = \|\nabla I_{MRI}\|^2$$

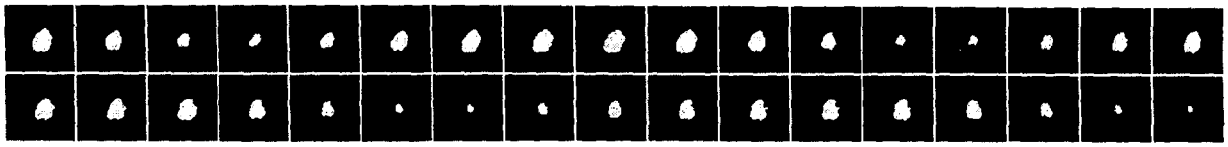

Figure 5. Training data: 2D binary shape models of the left ventricle based on human interactive segmentations of different spatial and temporal slices of a patient's cardiac MRI. Thirty four of the fifty training shapes are shown.

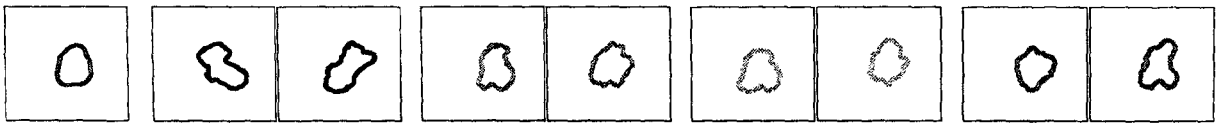

Figure 6. Shape variability of the left ventricle. (a) The mean shape. (b) & (c) $\pm 1\sigma$ variance of the 1^{st} principal mode. (d) & (e) $\pm 1\sigma$ variance of the 2^{nd} principal mode. (f) & (g) $\pm 1\sigma$ variance of the 3^{rd} principal mode. (h) & (i) $\pm 1\sigma$ variance of the 4^{th} principal mode.

where ∇ here denotes a 3D gradient operator. This mapping was employed because 1) the interior of the prostate is homogeneous in intensity, so with this mapping, the interior regions of the prostate are mapped to low values while the boundaries of the prostate are mapped to high values, and 2) this mapping is robust to the smooth spatially-varying intensity artifact cause by the ERC. We segment the prostate gland by applying the energy functional described in (4) on this transformed data set. We initialize the segmenting surface to be within the interior of the prostate gland. With each iteration, the segmenting surface moves outward to capture more and more of the low valued region in the transformed data which corresponds to the prostate gland. Eventually, the segmenting surface converges to a local minimum near the boundaries of the prostate corresponding to high values in the transformed data. Figure 10 shows the prostate segmentation results of two patients (neither of which are part of the training database) using $k = 5$. This figure, in addition to showing the 3D prostate gland, also shows the segmentation of the prostate in twelve contiguous 2D slices of the MR data set. The red curves in the figure show the segmentation by our algorithm and the green curves show the segmentation by a radiologist from Brigham and Women's Hospital (BWH). Another radiologist, also from BWH, rated the first radiologist's segmentation of patient A to be slightly better than the computer's, and rated the computer's segmentation of patient B to be slightly better than the radiologist's.

6 Conclusions

We have outlined a robust and efficient segmentation algorithm based on an implicit representation of the curve. This representation is set in an Eulerian framework hence does not require point correspondences. This approach is designed to delineate structures with shape variations consistent with the training data. It can be used to segment images contaminated by noise and delineate structures complicated by missing or diffuse edges. In addition, this framework is flexible, both in terms of its ability to model and segment complicated shapes, as well as its ability to accommodate the segmentation of multidimensional data sets. While we have not pursued it in this paper, our approach can also handle topological changes in the segmenting curve in a seamless fashion.

References

[1] A. Chakraborty, L. Staib, and J. Duncan, "An integrated approach to boundary finding in medical images," *IEEE Workshop on Biomedical Image Analysis*, pp. 13-22, 1994.

[2] T. Chan and L. Vese, "Active contours without edges," *IEEE Trans. on Image Process.*, to appear.

[3] T. Cootes, C. Taylor, D. Cooper, and J. Graham, "Active shape models–their training and application," *Comp. Vision and Image Understanding*, vol. 61, pp. 38-59, 1995.

[4] M. Leventon, E. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours," *IEEE Conf. on Comp. Vision and Patt. Recog.*, vol. 1, pp. 316-323, 2000.

[5] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Comm. Pure Appl. Math.*, vol. 42, pp. 577-685, 1989.

[6] S. Osher and J. Sethian, "Fronts propagation with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations," *J. of Comput. Phys.*, vol. 79, pp. 12–49, 1988.

[7] L. Staib and J. Duncan, "Boundary finding with parametrically deformable contour models," *IEEE Trans. Patt. Analysis and Mach. Intell.*, vol. 14, pp. 1061-1075, 1992.

[8] Y. Wang and L. Staib, "Boundary finding with correspondence using statistical shape models," *IEEE Conf. Comp. Vision and Patt. Recog.*, pp. 338-345, 1998.

[9] A. Yezzi, A. Tsai, and A. Willsky, "A statistical approach to snakes for bimodal and trimodal imagery," *Int'l Conf. on Comp. Vision*, vol. 2, pp. 898-903, 1999.

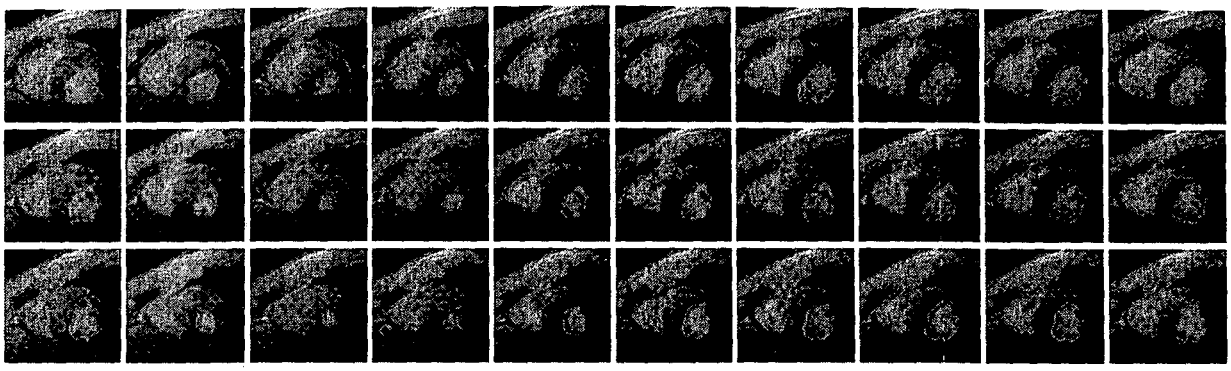

Figure 7. Left ventricle segmentation results of cardiac MRI. The segmentation by our algorithm (red curves) is compared to the segmentation by an interactive 4D cardiac MRI segmentation algorithm (green curves).

Figure 8. Training data: eight 3D shape models of the prostate gland obtained based on stacked 2D hand segmentations.

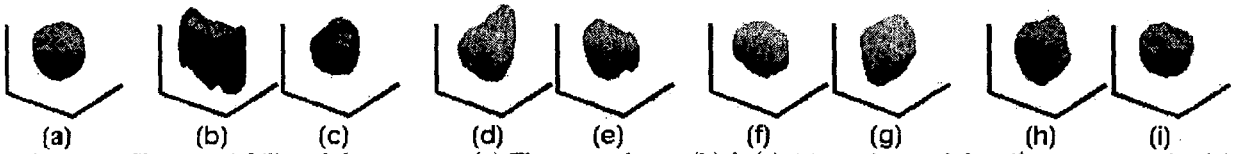

Figure 9. Shape variability of the prostate. (a) The mean shape. (b) & (c) $\pm 1\sigma$ variance of the 1^{st} principal mode. (d) & (e) $\pm 1\sigma$ variance of the 2^{nd} principal mode. (f) & (g) $\pm 1\sigma$ variance of the 3^{rd} principal mode. (h) & (i) $\pm 1\sigma$ variance of the 4^{th} principal mode.

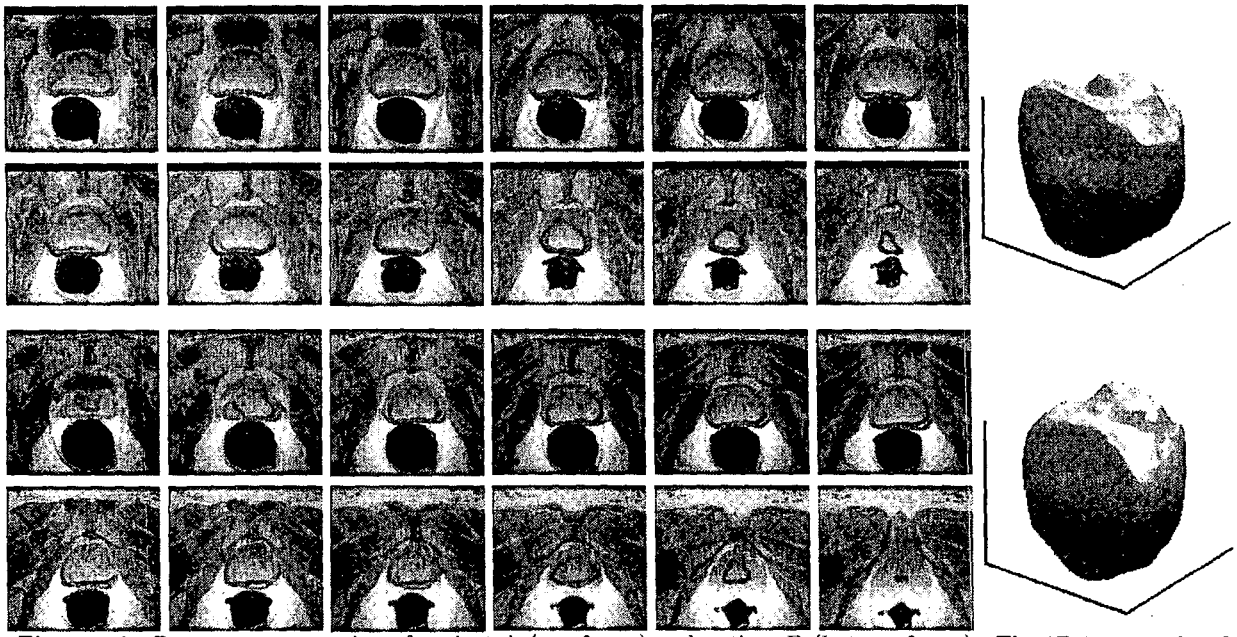

Figure 10. Prostate segmentation of patient A (top frame) and patient B (bottom frame). The 2D images of each patient represent consecutive axial slices of the prostate. The segmentation by the radiologist (green curves) is compared to the segmentation by our algorithm (red curves). The 3D model of the segmented prostate is shown on the right.